RICE UNIVERSITY

A characterization of the tail $\sigma$-field
for certain Markov chains

by

Cpt. Howard Floyd Bachman

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

Master of Arts

Thesis Director's signature:

Houston, Texas

May, 1972
ABSTRACT

A characterization of the tail $\sigma$-field for certain Markov chains

Cpt. Howard Floyd Bachman

If $C(\nu, P)$ is a countable state, recurrent, aperiodic and irreducible Markov Chain with stationary probabilities, then the measure of any set of the tail $\sigma$-field is equal to either zero or one. Although recurrent Markov chains have trivial tail $\sigma$-fields this is not in general true for transient chains. However the tail $\sigma$-field for two merging independent Markov chains is trivial. Investigating the invariant measurable sets of $\Omega$ leads to the solution of a functional equation of the general form $pf = f$. 
ACKNOWLEDGMENTS

The author wishes to express his deep thanks to Professor William A. Veech for his guidance in the form of his many helpful suggestions during the course of this work.
REFERENCES


0. Introduction

In 1955 Hewitt and Savage [9] presented a paper using a measure theoretical approach to certain probabilistic problems. From this paper has emerged a general approach to such problems which Blackwell and Freedman have used to explain the asymptotic behavior of certain Markov Chains.

Presented here is the description of the tail and invariant \( \sigma \)-fields for certain recurrent and transient Markov Chains. Complete identification of the characteristics of these two fields has made possible a better understanding of a number of diverse Markov Chain processes. Our first result is motivated by a paper of Blackwell and Freedman [2] in which they were able to prove that the tail \( \sigma \)-field of a large class of Markov Chains is trivial. More precisely, let \( \{X_n / 0 \leq n < \infty\} \) be a sequence of coordinate functions on the measure space \( (\Omega, \mathcal{B}, \mu) \). Define \( \Omega = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ldots \)

\[ B = \mathcal{B}(X_0, X_1, \ldots) \]

is the \( \sigma \) field of measurable sets on \( \Omega \) and \( \mu \) is a measure defined in Section 2. \( \mathcal{B}_n \) is the \( \sigma \)-field generated by \( \{X_n, X_{n+1}, \ldots\} \). Then define \( \mathcal{B}_\infty \), the tail \( \sigma \) field to be \( \bigcap_{n=0}^{\infty} \mathcal{B}_n \). Furthermore denote a Markov Chain with stationary probabilities \( P \) and an initial distribution function \( \nu \) by \( C(\nu, P) \).

**Theorem I.** If \( C(\nu, P) \) is a countable state, recurrent, aperiodic and irreducible Markov Chain with stationary probabilities, then the measure of any set of the tail
σ-field is equal to either zero or one.

While a recurrent Markov Chain has a trivial tail σ field, this is not, in general, true for transient chains. Although very little has been written in this field, one result which I feel is of interest deals with the tail field for each of two merging independent Markov Chains. Merging is defined in Section 3.

**Theorem II.** The tail σ field for two merging independent Markov Chains is trivial.

Finally Section 4 is devoted to an investigation of the invariant measurable sets of Ω. This effort will lead to the solution of a functional equation of the general form pf = f.

**Theorem III.** Let f(x), 0 < x < 1 be a Lebesgue integrable function such that either f is essentially bounded or f is non negative and suppose f(x) satisfies

\[
\frac{1}{2x} \int_{0}^{2x} f(y) dy \quad 0 < x \leq \frac{1}{2}
\]

\[
f(x) = \frac{1}{1-2(1-x)} \int_{1-2(1-x)}^{1} f(y) dy \quad \frac{1}{2} \leq x < 1
\]

then f(x) = A_0 + A_1 x for constants A_0, A_1.
1. Section one is devoted to certain preliminary notions which will be used throughout the paper.

1.1 Theorem: (Radon-Nikodym) Let \((\Omega, \mathcal{F}, \mu)\) be a \(\sigma\)-finite signed measure space, and let \(\nu\) be a \(\sigma\)-finite signed measure defined on \(\mathcal{F}\) which is absolutely continuous with respect to \(\mu\). Then there is an essentially unique measurable function \(f\) such that for every set \(E \in \mathcal{F}\),

\[\nu(E) = \int_E f \mu.\]

For proof see para. 23, p. 238 [12].

1.2 Definition: Let \((\Omega, \mathcal{F}, \mu)\) be a \(\sigma\)-finite signed measure space and let \(\mathcal{F}_0 \subseteq \mathcal{F}\) be a sub \(\sigma\)-field of \(\mathcal{F}\). Given a function \(f\) which is integrable with respect to \((\Omega, \mathcal{F}, \mu)\) one can define a measure \(\nu\) on \(\mathcal{F}_0\) by

\[\nu(E) = \int_E f \mu \quad (E \in \mathcal{F}_0)\]

\(\nu\) is clearly absolutely continuous with respect to \(\mu\). Thus by 1.1 there exists an essentially unique integrable function \(f_0\) for \((\Omega, \mathcal{F}_0, \mu)\) such that

\[\nu(E) = \int_E f_0 \mu \quad (E \in \mathcal{F}_0)\]

\(f_0\) is called the conditional expectation of \(f\) with respect to \(\mathcal{F}_0\).

1.3 Lemma: Let \(f\) be \((\Omega, \mathcal{F}, \mu)\) integrable and \(\mathcal{F}_0 \subseteq \mathcal{F}\) a sub \(\sigma\)-field. If \(g\) is \(\mathcal{F}_0\) measurable and \(g \circ f\) is integrable then
\[ E(g \circ f / \mathcal{F}_0) = g E(f / \mathcal{F}_0) \]

Proof: Due to the uniqueness of the Radon-Nikodym derivative, and the definition of conditional expectation, we need only show that

\[ \int g \circ f \mu = \int g E(f / \mathcal{F}_0) \mu \quad (E \in \mathcal{F}_0) \]

If \( g(\omega) = 0 \) or 1 depending if \( \omega \) is or is not a point of \( H \in \mathcal{F}_0 \), the equation becomes the obvious

\[ \int f \mu = \int E(f / \mathcal{F}_0) \quad (E \cap H \subseteq E \cap H) \]

Thus equation (1) is true immediately if \( g \) is a linear combination of functions of the type just considered. The general case follows using an approximation procedure based on the Lebesque Dominated Convergence Theorem. For example see Theorem B, pg. 85 [12].

1.4 Lemma: Let \( \mathcal{F}_1 \leq \mathcal{F}_0 \leq \mathcal{F} \) be \( \sigma \) fields, and \( g \) be \( (\Omega, \mathcal{F}, \mu) \) integrable. Then

\[ E(g / \mathcal{F}_1) = E(E(g / \mathcal{F}_0) / \mathcal{F}_1) \]

Proof: A consequence of the definition of condition expectation and the uniqueness of the Radon-Nikodym derivative.
1.5 Definition: The stochastic process formed by a sequence of integrable coordinate functions, \((X_0, X_1, X_2, \ldots)\) on \((\Omega, \mathcal{F}, \mu)\) is called a submartingale if

\[ E(X_{n+1} / \mathcal{F}_n) \geq X_n \quad \text{a.e.} \]

\(\mathcal{F}_n\) is defined as the \(\sigma\)-field generated by \((X_0, X_1, \ldots, X_n)\).

1.6 Lemma: If \((X_0, X_1, X_2, \ldots)\) is a submartingale and if \(\varphi\) is a real function which is monotone, non-decreasing and convex, with \(\varphi(X_n)\) integrable then \((\varphi(X_0), \varphi(X_1), \ldots)\) is a submartingale.

Proof: This is a straightforward consequence of Jensen's inequality. If \(\varphi\) be a convex function and \(X_n\) an integrable function on \((\Omega, \mathcal{F}, \mu)\), then

\[ E(\varphi(X_{n+1}) / \mathcal{F}_n) \geq \varphi(E(X_{n+1}) / \mathcal{F}_n) \]

1.7 Theorem: Let \(\{X_n, \mathcal{F}_n, n \geq 1\}\) be a submartingale. If

\[ \lim_{n \to \infty} E(|X_n| / \mathcal{F}_n) = H < \infty \]

then

(a) \(\lim_{n \to \infty} X_n = X_\infty\) exists with probability 1 a.e.

(b) \(E(|X_\infty| / \mathcal{F}_\infty) \leq H\)

Proof Pg. 324 [5].
1.8 Definition: Let $\mathcal{B}_n$ be the $\sigma$-field generated by $(X_n, X_{n+1}, \ldots)$. Then define $\mathcal{B}_\infty$, the tail $\sigma$-field to be
$$\bigcap_{n=0}^\infty \mathcal{B}_n.$$ 

1.9 Lemma: If $A \in \mathcal{B}_\infty$ then there exists an $A_n \in \mathcal{B}_n$ for each $n \in \mathbb{N}$ such that
$$A = \{x \in \mathbb{R}^\mathbb{N} : x_n \in A_n \}.$$

Proof: Define $A(y) = \{y_w : w \in A, y \in \mathbb{R} \}$ recall $w = (X_0, X_1, \ldots)$. Since $A \in \mathcal{B}_\infty$, this implies with respect to all $\mathcal{B}_n$, $n \geq 1$, that $A(X) = A(X') \forall X, X'$ since the elements of $\mathcal{B}_n$ depend only on the elements from the $n$th position on. Thus for $n = 0$ $A = \mathbb{R} \times A(X)$ i.e. $A_0 = A(X)$. By induction one can show this same result for any $n$. 
2. Section two

2.1 Definition: Let \( R \) be a non empty set and \( \mathfrak{F} \) a
\( \sigma \)-field of subsets of \( R \). Given a probability measure \( \nu \)
on \( \mathfrak{F} \), define the product measure

\[
\mu = \nu \times \nu \times \ldots \quad \text{on} \quad \Omega = R \times R \times \ldots
\]

(The Kolmogorov Theorem applies here, see 2.9).

2.2 Definition: A perutation, \( \pi \), is a one to one map of
the set of natural numbers onto itself. \( \pi \) has finite
support if \( \pi_n = n \) for all but a finite number of \( n \). Denote by \( \mathfrak{S} \) the set of all permutations with finite support. Associate to each \( \pi \in \mathfrak{S} \) a transformation
\( \tau_\pi : \Omega \to \Omega \) in the natural way. \( \tau_\pi \) is measure preserving for
the product measure \( \mu \). (p. 164, [2]).

2.3 Definition: A set \( B \in \mathfrak{B} \) is \( \mathfrak{S} \)-invariant if \( \tau_\pi^{-1} B = B \)
for all \( \pi \in \mathfrak{S} \).

2.4 Definition: A \( \mathfrak{B} \) measurable function, \( f \), is \( \mathfrak{S} \)-invariant
if \( f(\tau_\pi(w)) = f(w) \) for all \( \pi \in \mathfrak{S} \).

2.5 Theorem: (Hewitt-Savage) The product measure \( \mu \)
assumes the values zero or one on \( \mathfrak{S} \)-invariant sets that belong
to \( \mathfrak{B} \).
Proof: (Doob) Let $A \in \mathcal{B}$ be $\mathcal{B}$-invariant. Define

$$h(w) = \chi_A(w) - \mu(A) \quad (w \in \Omega)$$

and $\varphi(X_n) = E(h/X_n)$. Since $\mu$ is a product measure, one can write

$$\int \varphi(X_n) \varphi(X_m) \mu = c^2 \delta_{mn} \quad \Omega$$

($\delta_{mn}$ is the Kronecker delta).

$$C^2 = \int \varphi(X_1)^2 \mu \quad \Omega$$

note that $\int \varphi(X_m) = \int \varphi(X_n) = \int \varphi(X_1)$. Since $A$ is $\mathcal{B}$ invariant. Assume $C^2 \neq 0$. Then by Bessel's inequality (p. 152, [12]).

$$\int h^2 \mu \geq \sum_{n=1}^{N} \left[ \int \frac{h\varphi(X_n)}{C} \right]^2 = NC^2$$

But $h^2$ is bounded. Thus $C^2 = 0 \varphi(X_n) = 0$ a.e. By repeating the same argument, $E(h/X_1\ldots X_n) = 0$ a.e.

$$h = \lim_{n \to \infty} E(h/X_1\ldots X_n) = 0 \quad a.e.$$

But $h = \chi_A - \mu(A)$. This implies $\mu(A)$ is in the range of $\chi_A$. So $\mu(A) = 0$ or 1.
2.6 Definition: A matrix $P = \{P_{ij}\}$, $i,j \geq 1$ such that

(a) $P_{ij} > 0$ all $ij$

(b) $\sum_{j=1}^{\infty} P_{ij} = 1$ all $i$

is called a Markov transition matrix. Powers of $P$ are defined as follows:

$$P^{(n)}_{ij} = \begin{cases} 
\delta_{ij} & n = 0 \\
P_{ij} & n = 1 \\
\sum_{K=1}^{\infty} P^{(n-1)}_{iK} P_{Kj} & n \geq 2
\end{cases}$$

2.7 Lemma: $P^{(n)}$ forms another Markov transition matrix. (p. 9 [4]).

2.8 Definition $C(\nu, P)$ is a stationary Markov Chain with transition probabilities $P_{ij}$ and initial distribution $\nu$. $C(\nu, P)$ is a stochastic process, $\{X_n, \Omega, \mathcal{B}, \mu\}$ where $\nu$ is a non negative function on $\mathbb{R}$ such that $\sum_{j=1}^{\infty} \nu(j) = 1$ $j \in \mathbb{R}$.

Recall that a cylinder set has the following form.

$$C(A_0, \ldots, A_n) = \{X \in \Omega \cap X_0 \in A_0, \ldots, X_n \in A_n\}$$

Then define $\mu(C(A_0, A_1, \ldots, A_n)) = \sum_{j_1 \in A_1} \nu(j_1) P_{j_1, j_2} \cdots P_{j_{n-1}, j_n}$.
If \( v(i) = \delta_{ij} \) for some \( i \in R \) then we use special notation \( C(i,P) \) for \( C(v,P) \) and \( \mu_i \) for \( \mu \).

2.9 Remark. By Kolmogorov's Extension Theorem (p. 5 [13]) countable additivity holds on the cylinder field. Then by Caratheodory's Theorem (p. 17 [10]) it is always possible to extend a non-negative completely additive set function defined on a Borel Field to all sets of the \( \sigma \)-field without losing either of its properties. Furthermore this extension is unique.

2.10 Definition. A \( \mathcal{B} \) measurable function \( \tau(\omega) \), which is integer valued, is a stopping time for \( C(v,P) \) if for all \( n \)

\[
(a) \quad \{\omega / \tau(\omega) \leq n\} \in \mathcal{B}(X_0, X_1, \ldots, X_n)
\]

\[
(b) \quad \tau(\omega) < \infty \quad \text{a.e.}
\]

2.11 Remark. There are two natural measures that can be defined on \( \Omega \) given a stopping time \( \tau \).

Define a map \( T_\tau : \Omega \to \Omega \) by \( T_\tau(X_0, X_1, X_2, \ldots) = (X_\tau, X_{\tau+1}, \ldots) \). This is fine unless \( \tau(\omega) = \infty \). But by assumption this occurs on a set of measure zero, which can be ignored for our purposes. Using \( T_\tau \), we can set up the chain \( C(\nu_\tau, P) \).

\[
\nu_\tau(j) = \mu(\omega / X_\tau(\omega) = j)
\]
Denote $\mu_T$ as the measure defined in the natural way for $C(\nu_T, P)$. Another measure can be induced on $\Omega$ by $T_T$.

Define

$$\varphi_T(B) = \mu(T_T^{-1}B) \quad (B \in \mathcal{B})$$

2.12 Definition. If $\varphi_T(B) = \mu_T(B)$ then $C(\nu, P)$ is said to have the Strong Markov Property.

2.13 Theorem. $C(\nu, P)$ possesses the Strong Markov Property.

Proof: It suffices to verify $\mu_T(B) = \varphi_T(B)$ for cylinder sets $B$. We do this by induction.

$$\mu\{w/X_T \in A_1\} = \sum_{m=1}^{\infty} \mu\{w/\tau(w) = m, X_m \in A_1\}$$

For each $m$, the set $\{w/\tau(w) = m \text{ and } X_m \in A_1\}$ is $\mathcal{F}_1(X_1, \ldots, X_m)$ measurable. So $X(X_m \in A_1) = \sum_{j \in A_1} \mu\{w/\tau(w) = m, X_m = j\} = \sum_{j \in A_1} \nu_T(j)$. So the claim holds for $n = 1$. Assume it is true for $n - 1$.

$$\mu\{w/X_T \in \bigcap_{m=1}^{\infty} A_1, X_{\tau+n-1} \in A_n\} = \sum_{m=1}^{\infty} \mu\{w/\tau(w) = m, X_m \in A_1, \ldots, X_{m+n-1} \in A_n\}.$$

Then

$$\int_{X(X_m \in A_1, \ldots, X_{m+n-1} \in A_n)} = \sum_{j \in A_1} \mu\{w/\tau(w) = m, X_m = j\} \prod_{j \in A_1, j_n \in A_n} p_{j_1 j_2 \cdots j_{n-1} j_n}$$
\[ \begin{align*}
&= (\Sigma_{j_1 \in A_1} \cdots \Sigma_{j_n \in A_n} P_{j_1} P_{j_2} \cdots P_{j_{n-1}j_{n-1}} P_{j_n}) \sum_{j_n \in A_n} \mu\{w/X_{\tau_1} \in A_1, \ldots, X_{\tau_n} \in A_n\} P_{j_n-1j_n} = \mu\{w/X_{\tau_1} \in A_1 \ldots X_{\tau_n} \in A_n\}.
\end{align*} \]

2.14 Remark. The Strong Markov Property states that the process \( y_n(w) = X_{\tau(w)+n}(w) \) for \( \tau(w) \) a stopping time, forms another Markov Chain with the same transition probabilities. For a more complete discussion of this phenomenon see Chung pp. 168-177 [4].

2.15 Definition. Let \( r \in R \) be a fixed point. Then \( r \) is a recurrent state of \( \mathbb{C}(\gamma, P) \) if
\[ \mu_r\{w/X_n(w) = r \text{ for some } n > 1\} = 1. \]

2.16 Lemma. A point \( r \) is recurrent if and only if
\[ \sum_{n=0}^{\infty} P(n) = \infty. \] See pg. 22 [4].

2.17 Definition. Given a fixed \( i \in \mathbb{R} \) then the inter-\( i \)-blocks are defined as follows:

\[ \{(X_1, \ldots, X_n) | X_1 = i, x_j \neq i \ 2 \leq j \leq n, 1 \leq n \leq \infty\} \]
Define \( N_i(w) = \lim_{N \to \infty} \sum_{n=1}^{N} \delta_{X_n(w), i} \).

2.18 Construction. We can now construct the inter-i-block process which corresponds to a given Markov Chain. Define \( \Omega_i = \{ w/X_1 = i, N_i(w) = \infty \} \)

\[
R_i = R \cap \Omega_i \\
\land_i = R_i \times R_i \times \ldots
\]

Assuming \( c(i, P) \) is recurrent, then \( \mu_i(\Omega_i) = 1 \). Define the measure on \( \land_i \) by

\[
\eta(B) = \mu_i(T^{-1}B) \quad (B \in \land_i)
\]

\( T:\Omega_i \to \land_i \) such that \( T \) is the natural 1-1 map between an element of \( \Omega_i \) and the sequence of inter-i-blocks of \( \land_i \) which strung together forms the same element. Denote the coordinate functions of \( \land_i \) by \( \varphi_n \). Now define \( \gamma_n \) on the set of subsets of \( R_i \) by

\[
\gamma_n(E) = \mu_i\{ w/\varphi_n(T(w)) \in E \} = \mu_i[T^{-1}\varphi_n^{-1}E]
\]

Let \( \tau_n(w), w \in \Omega_i \), be the index \( t \) for which \( X_t(w) = i \) for the \( n \)th time. Since \( i \) is recurrent \( \tau_n \) is finite on \( \Omega_i \). By the Strong Markov property,
\[ \gamma_n(E) = \mu_1(\omega/\nu_n(T(\omega) \in E) \]

\[ = \mu_1(\omega/\nu_1(T,\omega) \in E) = \gamma_1(E) \]

Additionally, we claim \( \eta \) is the product \( \eta = \gamma \times \gamma \times \ldots \). This implies,

\[ \eta[\lambda \in \land \varphi_1(\lambda) \in B_1, \ldots, \varphi_n(\lambda) \in B_n] \]

\[ = \gamma(B_1) \gamma(B_2) \ldots \gamma(B_n) \quad \text{for all sets} \]

\( B_1, \ldots, B_n \leq \mathcal{A} \). It is enough to verify when all \( \varnothing_i \) are singleton sets. Let \( B_1 = \{a\}, B_2 = \{a_2\}, B_n = \{a_n\} \). If

\[ a_{\ell} = (y_1(\ell), y_2(\ell), \ldots, y_n(\ell)) \]

\[ = (i, y_2(\ell), \ldots, y_n(\ell)) \quad \ell = 1, 2, \ldots, n \]

the definition of \( \eta \) in terms of \( \mu_1 \) leads us to

\[ \eta[\lambda \in \land \varphi_1(\lambda) = a_1, \ldots, \varphi_n(\lambda) = a_n] \]

\[ = \prod_{\ell=1}^{n} \prod_{j=1}^{n-1} P(y_j(\ell) \mid y_{j+1}(\ell)) \]

\[ = \eta[\varphi_1(\lambda) = a_1] \eta[\varphi_2(\lambda) = a_2] \ldots \eta[\varphi_n(\lambda) = a_n] . \]

Thus the process \( \{\land, \eta, \varnothing(\varphi_1, \varphi_2, \ldots), [\varphi_n]\} \) which consist of independent, identically distributed random variables is the inter-i-block process associated with our given Markov chain.
2.19 Theorem. If $i$ is a recurrent state, $C(i, P)$ obeys the zero-one law.

Proof [2] Let $B \in \mathfrak{G}_\infty$. Claim $T(B) = \tilde{B} \in \mathfrak{G}(\Phi_1, \Phi_2, \ldots)$ is $\mathfrak{G}$-invariant. Let $\pi \in \mathfrak{G}$ such that $T_\pi(\lambda) \in \tilde{B}$ for some $\lambda$. If $T(\omega) = \lambda$ for some $\omega$, then there is some permutation $\pi_\omega \in \mathfrak{G}$ such that $T(T_\pi(\omega)) = T_\pi(\lambda)$. Therefore $T_\pi(\omega) \in \tilde{B}$ which implies $\omega \in \tilde{B}$. But then $\lambda \in \tilde{B}$ and $T_\pi^{-1} \tilde{B} \subseteq \tilde{B}$. But $T_\pi^{-1} = T_\pi^{-1}$. So $T_\pi^{-1} \tilde{B} = \tilde{B}$. Then $\tilde{B}$ is $\mathfrak{G}$ invariant and therefore by 2.5, $\eta(\tilde{B}) = 0$ or 1.

2.20 Definition. The matrix $P$ is irreducible if for each pair of states, $i$ and $j$, there exists $n > 0$ such that $P_{ij}^{(n)} > 0$.

2.21 Lemma. Given an irreducible Markov Transition matrix, all states are either recurrent or non-recurrent PROOF p. 31 [4].

2.22 Definition. An irreducible Markov transition matrix is aperiodic if $d(i) = 1$

$$d(i) = \gcd \Gamma_i$$

$$\Gamma_i = \{n \geq 1 | p_{ii}^{(n)} > 0\}$$

2.23 Theorem: (Blackwell-Friedman) Let $C(v, P)$ be a Markov Chain in which $P$ is irreducible, aperiodic, and
all states are recurrent. Then \( C(v, P) \) obeys the zero-one law.

Proof: Let \( B \in B_\infty \). There exists a \( B_m \in \cap_{n=m+1}^\infty B(X_n, X_{n+1}, \ldots) \) such that

\[
\begin{align*}
B &= R \times R \times \ldots \times R \times B_m \\
\mu_i(B) &= \sum_{j=1}^{\infty} P_{ij}^{(m)} \mu_j(B_m)
\end{align*}
\]

By 2.19 \( \mu_j(B_m) = 0 \) or 1 for all \( j \) and \( m \). Also \( \mu_i(B) = 0 \) or 1. Since \( \sum_{j=1}^{\infty} P_{ij}^{(m)} = 1 \), \( \mu_i(B) = \mu_j(B_m) \) for all \( j, m \) such that \( P_{ij}^{(m)} > 0 \). Since \( P \) is irreducible and aperiodic, there exists an \( m \), for each \( j \), such that \( P_{ij}^{m} > 0 \) for all \( m' \geq m \). Thus the limit of \( \mu_j(B_m) \) exists and is zero or one for all \( j \), the same alternative for each \( j \). Clearly \( \mu_i(B) \) is determined by \( \lim_{m \to \infty} \mu_j(B_m) \).

2.24 Lemma. Let \( X_1, X_2, \ldots \) be a stochastic process. \( \mathfrak{G}_n = \mathfrak{G}(X_n, X_{n+1}, \ldots) \). Then the zero-one law holds iff

\[
\lim_{n \to \infty} \sup_{B \in B_n} \left| \mu(A \cap B) - \mu(A) \mu(B) \right| = 0
\]

\( \forall A \in B \)

Proof. By the Martingale Theorem p. 115 [1]
\[
\lim_{n \to \infty} \|E(\chi_A | B_n)\|_1 = \|E(\chi_A | \mathcal{B}_\infty)\|_1
\]

and \(E(\chi_A | \mathcal{B}_\infty) = \mu(A)\) a.e. If \(B \in \mathcal{C}_n\), then \(\mu(A \cap B)\)
\[
= \int E(\chi_A | \mathcal{B}_n) \mu + \sup_{B \in \mathcal{C}_n} \left| \int E(\chi_A | \mathcal{B}_n) \mu - \mu(A) \mu(B) \right|
\]
\[
- \sup_{B \in \mathcal{C}_n} \left| \int E(\chi_A | \mathcal{B}_n) - \mu(A) \right| \mu \leq \|\int E(\chi_A | B_n) - \mu(A)\| \mu \to 0
\]

by the assumed 0 or 1 law. Conversely,

If \(\lim \sup_{n \to \infty} \left| \mu(A \cap B) - \mu(A) \mu(B) \right| = 0\) then
\[
\lim \sup_{n \to \infty} \mu(B) = 0 \text{ or } 1 \text{ but } B_n \to B \text{ which implies } B \in \mathcal{B}_\infty.
\]

Thus the zero-one law holds.

2.25 Theorem (Orey ) Let \(P\) be the transition matrix for an irreducible, aperiodic, recurrent Markov Chain.

For all states \(i, k\)
\[
\lim_{n \to \infty} \sum_{j=1}^{\infty} |p_{ij}^{(n)} - p_{kj}^{(n)}| = 0
\]

Proof [7]: Define \(\nu\), the probability measure on \(\mathbb{R}\).
\(\nu(j) = \frac{1}{2}(\delta_{ij} + \delta_{kj})\), Let \(A, B \in \mathcal{B}\) such that
A = \{ w/ X_1(w) = i \} \\
B = \{ w/ X_n(w) \in K \}

K is a non empty subset of R

Now \( \mu(A \cap B) = \frac{1}{2} \sum_{j \in K} P_{ij} \), \( \mu(A) \mu(B) = \frac{1}{2} \sum_{j \in K} P_{ij} + P_{kj} \) 

so \( |\mu(A \cap B) - \mu(A) \mu(B)| = \frac{1}{2} |\sum P_{ij} - P_{kj}| \). Since

\( \sum P_{ij} = 1 = \sum P_{kj} \) then

\[
\sum_{j=1}^{\infty} |P_{ij} - P_{kj}| = 2 \sup_{S \subseteq K} |\sum_{j \in K} P_{ij} - P_{kj}|. 
\]

Let \( K = \{ j | P_{ij} - P_{kj} \geq 0 \} \) and so

\[
\sum_{j=1}^{\infty} |P_{ij} - P_{kj}| \leq \sup_{B \in \mathcal{C}_n} |\mu(A \cap B) - \mu(A) \mu(B)|. 
\]

The result follows from 2.25.
3.1 Definition. Assume \( (\Omega, \Theta, \mu) \) are the same as in the previous section. The same stochastic process \( (\nu, P) \) will be considered. Define \( T: \Omega \rightarrow \Omega \) such that \( T(X_n) = X_{n+1} \). Let \( R(n,i) = \{ j \in R \mid P_{ij}^n > 0 \} \).

3.2 Definition. A state \( i \in R \) is merging if two independent Markov Chains with the same initial state space \( R \), transition probabilities \( P_{jk} \) and initial state \( i \) meet infinitely often with probability 1.

3.3 Remark. Note that the above definition states that the chains meet but does not say that they return to state \( i \) infinitely often.

3.4 Lemma. Let the state \( i \) be merging. Suppose \( h, g \in R(n,i) \). Then two independent Markov Chains with state space \( R \), stationary transition probabilities \( P_{jk} \) and initial states \( g \) and \( h \) respectively meet infinitely often with probability 1.

Proof. Clearly since two chain which start at \( i \) have a positive probability of being in \( g, h \) respectively at time \( n \) and since these chains must still meet infinitely often, it must follow that the two chains starting from \( g \) and \( h \) must meet infinitely often.

3.5 Lemma. Suppose \( i \in R \) is merging and \( g, h \in R(n,i) \). Then \( \mu_g(A) = \mu_h(A) \forall A \in \mathcal{B}_\infty \).
Proof. Let \( \{X_t / t \geq 0\} \), \( \{y_t / t \geq 0\} \) be independent Markov Chains with initial states \( g \neq h \). \( X_t, y_t \) are coordinate functions. Define a new chain \( \{z_t / t \geq 0\} \) by the following:

\[
X_t \text{ provided } X_i \neq y_i \quad 1 \leq i \leq t,
\]

\[
z_t = y_t \quad \text{otherwise}
\]

\( \{z_t / t \geq 0\} \) is initially distributed like \( \{X_t / t \geq 0\} \). Using our measure \( \mu \) and \( \mu_1 \) as defined in section 2.8, for \( A \in \mathfrak{B}_\infty \) by 3.4

\[
\mu_1(A) = \mu([X_t / t \geq 0] \cap A) \\
= \mu([x_t / t \geq 0] \cap A) \\
= \mu([y_t / t \geq 0] \cap A) \\
= \mu_h(A)
\]

3.6 Theorem. If \( i \in R \) is merging then \( \mu_1 = 0 \) or 1 on \( \mathfrak{B}_\infty \).

Proof. Let \( A \in \mathfrak{B}_\infty \) and \( w, w' \) be elements of each chain

\[
E(\chi A / X_0, \ldots X_n)(w) = \mu_{X_n}(w) (T^{-n}A) \quad (\text{a.e. } \mu_1)
\]

Chain \( X_n(w) = X_n(w') \) a.e. since the two chains are merging. Then \( E(\chi A / X_0, \ldots X_n)(w) = E(\chi A / (X_0, \ldots X_n)(w')). \) Thus by 3.5 \( E(\chi A / X_0, \ldots X_n) \) is constant a.e. \( \mu_1 \). As \( n \to \infty \)
E(χA/X_0, ..., X_n) \rightarrow χA (a.e. μ_1) by the martingale theorem. Since χA is constant (a.e. μ_1) then μ_1(A) = 0 or 1.
4.0. In this section we will deal with the problem of finding the solutions to the equations of the general form \( pf = f \).

4.1 Definition. Let \((\Omega, \mathcal{B}, \mu)\) be a measure space with \(\Omega\) the countable product space of the unit interval \(I\). In the next few paragraphs we will develop the definition of the measure \(\mu\).

4.2 Definition. Let \( F_n : \Omega \to I^{n+1} \) be defined as

\[
F_n(\omega) = (X_0(\omega), X_1(\omega), \ldots, X_n(\omega)) \quad \omega \in \Omega
\]

Assume \(\mu\) is defined and let \( \lambda_n(B) = \mu(F_n^{-1}B) \) (B, borel).

4.3 Definition. Let \( P(x,y) \) be a Borel function on \(I \times I\) such that

1. \( P(x,y) \geq 0 \quad \forall x, y \in I \)
2. \( \int_I P(x,y) dy = 1 \)

4.4 Definition. Let \( \alpha(X_0, \ldots, X_{n+1}) \) be an integrable function with respect to \( \lambda_n \) which depends only on the last coordinate. Let \( \mathcal{C}_n \) be a collection of Borel sets which depend on the first \( n+2 \) coordinates. Then

\[
E(\alpha/\mathcal{C}_n) = \int_{-\infty}^{\infty} P(x_n, y) \alpha(y) dy.
\]

By induction for \( A_0, A_1, \ldots, A_n \)
borel sets in $I$

$$\lambda_n\{\tilde{A}_0 \cap A_1 \cap \ldots \cap \tilde{A}_n\} = \int \ldots \int p(X_0, X_1) \cdots p(X_{n-1}, X_n) \delta X_0 \ldots dX_n$$

where $\tilde{A}_j = \{X \in \mathbb{I}^{j+1} | X_j \in A_j\}$.

4.5 Definition. Our desired measure $\mu$ is defined on the cylinder sets as follows

$$\mu[C(A_0, \ldots, A_n)] = \lambda_n[C(A_0, \ldots, A_n)] .$$

A standard result shows that this measure is consistent and so by Kolmogorov's Extension Theorem $\mu$ extends to a measure on $\mathcal{B}$.

4.6 Definition. Let $x$ be any real number such that $0 < x < 1$. Let $\delta(x)$ be an arbitrary measurable function on $I$ such that:

a) $0 < \delta(x) \leq \min(x, 1-x)$

b) $\inf_{x \in K} \delta(x) > 0$ for all compact sets $K$

Define

$$P(x, y) = \frac{1}{2\delta(x)} 0 < x < 1 x - \delta(x) < y < x + \delta(x) .$$

0
where \( I(x) = (x - \delta(x), x + \delta(x)) \). Thus \( \int P(x, y) = 1 \).

4.7 Remark. If \( E \) is any Borel set in \( I \), then

\[
\int P(x, y) \, dy = \int \frac{1}{2\delta(x)} m = \frac{m(E \cap I(x))}{\text{length } I(x)}
\]

\( m \) is the Lebesgue measure. Thus \( P(x, y) \) is a measurable density function of a probability measure on \( E \), that reflects the process of choosing a \( y \) at random from \( I(x) \).

4.8 Lemma. \( E(\|X_{n+1} - X_n\| \leq \delta_n) = \frac{1}{2} \delta(X_n) \) a.e.

Proof. Let \( A_0, \ldots, A_n \) be Borel sets

\[
\int |X_{n+1} - X_n| \, \mu = \int |X_{n+1} - X_n| \lambda_{n+1}
\]

\[
= \int \int_0^1 |X_{n+1} - X_n| \, P(X_n, X_{n+1}) \, \lambda_n
\]

\[
= \int \frac{1}{2} \delta(X_n) \, \lambda_n
\]

4.9 Lemma. The sequence \( X_0, X_1, \ldots \) converges both pointwise a.e. and in the \( L_1 \) norm to a limit function \( X_\infty \) which assumes the values 0 or 1 a.e. \( \mu(X_\infty(w) = 0 \) our starting point.

Proof. Since \( \{X_n\} \) is a uniformly bounded martingale, by
1.7, the sequence converges pointwise and in norm. Define 
\[ \|X_n\| = \min \{X_n, 1-X_n\} \]. We will show in the next remark 
\[ \int \|X_n\| \mu \leq 2\|X_{n+1} - X_n\|_1 \]. 

If you assume this is true, then \( \|X_n\| \rightarrow 0 \) in measure.
This implies \( X_\infty = 0 \) or 1 a.e. Now

\[ X_0(\omega) = \int X_0 \mu = \int X_\infty \mu = E(X_\infty) = 1 \cdot \mu \{X_\infty = 1\} \]

4.10 Remark. If we use \( \| \cdot \| \) as defined in Lemma 4.9 as our distance function \( \frac{1}{2} \delta(X) \) in Lemma 4.8 then since 
\[ \|E(f|B^n)\|_1 \leq \|f\|_1 \]. Apply 4.8 to obtain

\[ \|(\|X_n\|)\|_1 = 2\|E(X_{n+1} - X_n)\|_{B^n}\|_1 \leq 2\|X_{n+1} - X_n\|_1 \].

4.11 Definition. \( U \) is a Borel measurable subset of \( I \).
Define \( R(U) = \lim \sup_{X \rightarrow 1} \frac{m(U \cap (X, 1))}{1-X} \). \( R(U) \) is called the \( X \rightarrow 1 \) upper density of \( U \) at 1.

4.12 Lemma. If \( R(U) > 0 \) then \( \{X_0, X_1, \ldots\} \) visits \( U \) infinitely often with probability \( X_0 \).

Proof. We let \( \delta(X) = \|X\| \) for the rest of this discussion.
Define \( U_0 \) to be the union of \( U \) and its reflection about \( X = \frac{1}{2} \). If \( X_n \) visits \( U_0 \) infinitely often with probability
X_0, this will imply the lemma since almost all paths which converge to 1 must then visit U infinitely often.

Assume R(U) = \alpha > \beta > 0. Choose

0 < s_1 < s_2 < s_3 ... < s_n ... \leq 1

such that

a) \ m \{(s_n, 1-nU] > \beta(1-s_n)\}

b) \sum_{n=1}^{\infty} \frac{1-s_n+1}{1-s_n} < \infty

Define \( V_n = (0, 1-s_n) \cup (s_n, 1) \) and \( \tau_n: \Omega \to R \) such that

\( \tau_n(\omega) = \min j \) so \( X_j(\omega) \in V_n \)

By 4.9 \( X_\infty = 0 \) or 1 for almost all \( \omega \) implies \( \tau_n(\omega) < \infty \) for almost all \( \omega \). (i.e. \( \tau_n(\omega) \) is a stopping time) Assume \( s_n \geq 3/4 \) and let the set \( B(Borel) \subset I \) be symmetric about the point 1/2. We want to estimate the measure of the set where \( X_\tau \in B \). Assume \( \tau_n \neq 0 \) is \( X_0 \notin V_n \). Define

\( \mu_j = \mu \{ w/\tau_n(\omega) = j \) and \( X_j(\omega) \in B \} \)

This probability can be evaluated using our functions \( P(x,y) \).
\[ \mu_j \{ C(V_n^C, \ldots, V_n^C, B) = \int \ldots \int \frac{P(X_0, X_1) \ldots P(X_{j-1}, X_j)}{\lambda_{V_n^C}} \, dX_1 \ldots dX_j \}
\]

If \( 0 < X_{j-1} < 1 \) the symmetry of \( B \) implies

\[ \frac{1}{2} \frac{m(V_n \cap B)}{2 \|X_{j-1}\|} \leq \int \frac{P(X_{j-1}, X_j)}{V_n \cap B} dX_j \leq \]

\[ \left\{ \begin{array}{l}
\frac{m(V_n \cap B)}{2 \|X_{j-1}\|} & \text{if } |S_{j-1} - \frac{1}{2}| \geq \frac{1-n}{2} \\
\frac{m(V_n \cap B)}{2 \|X_{j-1}\|} + \frac{m(V_n)}{4 \|X_{j-1}\|} & \text{if } |X_{j-1} - \frac{1}{2}| \leq \frac{1-s_n}{2} 
\end{array} \right. \]

For \( s_n \geq 3/4 \), \( \|X_{j-1}\| \geq 3/8 \) for \(|X_{j-1} - \frac{1}{2}| \leq \frac{1-s_n}{2} \)

\[ \frac{2}{3} m(V) = \frac{m(V)}{4 \|X_{j-1}\|} \cdot \text{Define } I_j = \int \ldots \int \frac{P(X_0, X_1) \ldots P(X_{j-1}, X_{j-1})}{2 \|X_{j-1}\|} \, dX_1 \ldots dX_{j-1} \]

recalling (*)

\[ \frac{m(V_n \cap B)}{2} I_j \leq \mu_j \{ w/\tau_n(w) = j, X_j(w) \in B \} \]

Assuming \( X_0 \) such that \( |X_0 - \frac{1}{2}| > \frac{1-s_n}{2} \), \( j \geq 2 \). Then

\[ \mu_j \{ w/\tau_n(w) = j, X_j(w) \in B \} \leq \frac{m(V_n \cap B)}{2} I_j + \frac{3}{2} \, m(V_n) K_j \]
\[ K_j = \int \ldots \int P(X_0, X_1) \ldots P(X_{j-2}, X_{j-1}) \, dx_1 \ldots dx_{j-1} \]

\[ c_n = \{ x \mid x - \frac{1}{2} \leq \frac{1-S_n}{2} \} \]

\[ \int_{c_n} P(X_{j-2}, X_{j-1}) \, dx_{j-1} \leq \begin{cases} \frac{1-S_n}{2\|x_{j-2}\|} & \|x_{j-2}\| \geq \frac{S_n}{4} \\ 0 & \|x_{j-2}\| \leq \frac{S_n}{2} \end{cases} \]

Since the integral is non-negative implies

\[ 2\|x_{j-2}\| \geq \frac{S_n}{2} \geq \frac{3}{8} \]

\[ \int_{c_n} P(X_{j-2}, X_{j-1}) \, dx_{j-1} \leq \frac{4}{3} m(V_n) \quad (1-S_n = \frac{m(V_n)}{2}) \]

\[ K_j \leq \frac{4}{3} m(V_n) \int \ldots \int P(X_0, X_1) \ldots P(X_{j-3}, X_{j-2}) \, dx_1 \ldots dx_{j-2} \]

If \( \int V_n P(x, y) \, dy \geq (1-S_n) \Rightarrow \int V_n P(x, y) \, dy \leq S_n \) for \( j \geq 2 \)

\[ \mu_j \leq \frac{m(B \cap V_n)}{2} I_j + 2m(V_n)(S_n)^{j-2} \]

Define a new measure \( \nu_j = \mu\{w/\tau_n(w) = j\} \quad \frac{m(V_n)}{2} I_j \leq \nu_j \)
\[ \mu_j \leq \begin{cases} \frac{m(B \cap V_n)}{m(V_n)} & \nu_1 \\ \frac{m(B \cap V_n)}{m(V_n)} \nu_j + 2m(V_m)^2 + (S_n)^{j-2} & j \geq 2 \end{cases} \]

Then \( \sum_{j=1}^{\infty} \mu_j \leq \frac{m(B \cap V_n)}{m(V_n)} + 4m(V_n) \). Denote \( M(n,k) = \mu\{w / f_{n+j}(w) \in B_{n+j} \} \)

\[ 0 \leq S \leq k \} \] where \( B_n, B_{n+1} \ldots \) are symmetric Borel sets in \( V_n, V_{n+1}, \ldots \) such that \( B_{n+S} \subseteq V_{n+j} \cap V_{n+S+1} \). Using an inductive argument one can bound the measure of \( M(n,k) \)

\[ \sum_{t=1}^{\infty} \int \ldots \int \frac{P(X_{j+1}, y_1) \ldots P(y_{t-1}, y_t) dy_1 \ldots dy_t}{V_n C_n B_{n+1}} \leq \frac{m(B_{n+k} \cap V_{n+k})}{m(V_{n+k})} + 4m(V_{n+k}) \leq \prod_{S=0}^{k} \frac{m(B_{n+S} \cap V_{n+S})}{m(V_{n+S})} + 4m(V_{n+j}) \]

\[ \frac{m(B_{n+S})}{m(V_{n+S})} \leq 1 - \lambda < 1 \]

for \( n \) large

\[ 4m(V_{n+S}) \leq \frac{\lambda}{2} \] for all \( j \)

\[ \therefore M(n,k) \leq (1 - \frac{\lambda}{2})^K \rightarrow 0 \] as \( K \rightarrow \lambda \).
Given $\epsilon > 0$ choose $n_0$ such that

$$\mu\{\omega/\tau_{n_0}(\omega) < \tau_{n+1}(\omega) < \ldots\} > 1 - \epsilon$$

choose $k$ such that $(1 - \frac{1}{2})^k < \epsilon$. Then for any $n \geq n_0$

$$m\{\omega/\tau_n(\omega), \ldots \tau_{n+k}(\omega) \notin U_0\} \leq 2\epsilon$$

Thus, the set of $\omega$ such that infinitely many of the $X_n(\omega) \in U_0$ has measure at least $1 - 2\epsilon$. Since $\epsilon$ is arbitrary, the lemma is proved.

4.13 Theorem III. Let $f(X), 0 < X < 1$ be a Lebesgue integrable function such that either

a) $f$ is essentially bounded

b) $f$ is non negative

Let $f(X) = \begin{cases} \frac{1}{2X} \int^{2X} f(y)dy & 0 < X \leq \frac{1}{2} \\ \frac{1}{2(1-X)} \int^1 f(y)dy & \frac{1}{2} \leq X < 1 \end{cases}$

Then $f(X) = A_0 + A_1 X$ for constants $A_0, A_1$.

Proof. Assume condition (a). Clearly $f$ is continuous.
Define $g_n(w) = f(X_n(w))$, $n = 0, 1, ..., w \in \Omega$. Then

\{g_n(w), \beta^n\} forms a bounded martingale. Thus $g_n \to g_\infty$

and $E(g_n) \to E(g_\infty)$ (see 2.3). Define

$$\Omega_i = \{w/\lim_{n \to \infty} X_n(w) = i\} \quad i = 0, 1, ...$$

It is enough to show $g_\infty$ is constant a.e. on $\Omega_i$. Additionally

if these constants depend only on $f$ and not $X_0$, the

starting point, then $f(X_0) = E(g_0) = E(g_\infty) = X_0 A_1 + (1-X_0)A_0$.

Define $U_M, L_M \leq (0,1) \quad M \in \mathbb{R}$ by

$$U_M = \{X/f(X) \geq M\} \cap (\frac{1}{2}, 1)$$

$$L_M = \{X/f(X) < M\} \cap (\frac{1}{2}, 1)$$

Let $M_1 = \text{lub}\{M/R(U_M) > 0\}$ ($R(U_M)$ is upper density). Claim

g_\infty = M_1 \text{ a.e. on } \Omega_i$. Given $\epsilon > 0$

$$R(U_{M_1} + \epsilon) = 0 \quad R(L_{M_1} + \epsilon) = 1$$

Additionally $R(U_{M_1} - \epsilon) > 0$. As a result of (4.3) $X_n(w)$

visits both $L_{M_1} + \epsilon$ and $U_{M_1} - \epsilon$ infinitely often. It

follows immediately that $g_\infty = M_1 \text{ a.e. on } \Omega_i$. By a

similar argument one can show that $g_\infty = M_2 \text{ a.e. on } \Omega_i$.

($M_2 = \text{glb}\{M/R(L_M) > 0\}$). Thus the result follows.
4.14 Lemma. If \( f \) as defined in 4.13 is non negative, then it is essentially bounded.

Proof. Suppose \( \frac{1}{2} < X' < X < 1 \) and \( f(X) < C \). Since \( f \) is non negative

\[
f(X') = \frac{1}{2(1-X')} \int_{1-2(1-X')}^{1} f(y) dy \leq \frac{1}{2(1-X)} \int_{1-2(1-X')}^{1} f(y) dy \]

\[
= \frac{1-X}{1-X'} f(X) > \frac{1-X}{1-X'} C
\]

If \( \frac{1-X}{1-X'} > \frac{1}{3} \) then \( f(X') > \frac{C}{3} \). The range of \( X' \) is \( 1-3(1-X) < X' < X \). Thus if \( \limsup f(X) > C \) then \( R(U_{\frac{C}{3}}) > \frac{1}{2} \)

\( M_1 \geq \frac{C}{3} \). Since a martingale converges a.e. to a finite, \( M_1 < \infty \), then \( \frac{C}{3} \leq M_1 < \infty \) and \( f \) is bounded.